

## **REPORT DOCUMENTATION PAGE**

AFRL-SR-BL-TR-01-

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate of Davis Highway, Suite 1204, Arlington, VA 22202-4320. Respondents should be aware that notwithstanding any other provision of law, no person is required to respond to a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RET

Comply with a collection of information if it does not display a currency valid OMB control number. PLEASE DO NOT RETAIN				
<b>1. REPORT DATE (DD-MM-YYYY)</b> 31-03-2001		<b>2. REPORT TYPE</b> Final		
<b>3. DATES COVERED (From - To)</b> 01-07-1997 to 31-12-2000				
<b>4. TITLE AND SUBTITLE</b> Scalable Software for Quantum Molecular Dynamics				
<b>5a. CONTRACT NUMBER</b>				
<b>5b. GRANT NUMBER</b> F49620-97-1-0505				
<b>5c. PROGRAM ELEMENT NUMBER</b>				
<b>6. AUTHOR(S)</b> Gregory A. Voth				
<b>5d. PROJECT NUMBER</b>				
<b>5e. TASK NUMBER</b>				
<b>5f. WORK UNIT NUMBER</b>				
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>				
Department of Chemistry 315 S. 1400 E. Rm 2020 University of Utah Salt Lake City, UT 84112				
<b>8. PERFORMING ORGANIZATION REPORT</b>				
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>				
Air Force Office of Scientific Research/NL 801 N. Randolph St., Rm Arlington, VA 22203-1977				
<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b> AFOSR/NL				
<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>				
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>		AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR) NOTICE OF TRANSMITTAL DTIC. THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLIC RELEASE LAW AFR 100-12. DISTRIBUTION IS UNLIMITED.		
<b>13. SUPPLEMENTARY NOTES</b>				
<b>14. ABSTRACT</b>				
The objective of this project was to develop highly scalable quantum molecular dynamics software for the DoD CHSSI Initiative. Parallel versions of path integral Molecular Dynamics (PIMD) and centroid Molecular Dynamics (CMD) codes for homogeneous systems such as cryogenic solid hydrogen HEDM were completed. A code for multi-site homogeneous molecular systems was also finished, as well as a code for multi-site inhomogeneous systems. The parallel scaling of the codes was established on the IBM SP and SGI O2000 systems, as well as Linux clusters, over 128 processors.				
<b>15. SUBJECT TERMS</b>				
<b>16. SECURITY CLASSIFICATION OF:</b>		<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>
a. REPORT	b. ABSTRACT	c. THIS PAGE	uu	6
				<b>19b. TELEPHONE NUMBER (include area code)</b> 801-581-7272

**Standard Form 298 (Rev. 8-98)**

20010406 178

# AFOSR Final Technical Report

*Scalable Software for Quantum Molecular Dynamics*

(Grant #F49620-97-1-0505)

Principal Investigator: Gregory A. Voth  
Department of Chemistry  
University of Utah  
Salt Lake City, Utah 84112

## Objectives:

The objective of this project was to develop highly optimized quantum molecular dynamics software for the DoD CHSSI Initiative. This grant was initiated July 1, 1997. Progress during the duration of the grant is summarized below.

Traditional MD algorithms can be readily extended to incorporate equilibrium quantum mechanical effects through the use of discretized Feynman path integrals [1]. The resulting algorithm, path integral MD (PIMD), is highly amenable to parallelization because each physical particle becomes rigorously mapped onto a collection of classical-like quasiparticles at different values of imaginary time, each connected to its neighbor by a harmonic-like force. The discretized PIMD algorithm can be load-balanced across nodes because the quasiparticles with the same imaginary time index experience identical forces. The inter-node communications are minimal since there are only nearest-neighbor harmonic forces to compute. The basic PIMD algorithm, no matter how complex the physical system, can always be made to achieve excellent scalability and thus high performance. Moreover, more advanced classical MD techniques such as constant temperature and pressure algorithms can be readily incorporated into the PIMD algorithm. The PIMD algorithm is even amenable to a two-tier level of parallelism, the first being over the imaginary time slices (quasiparticles), the second over the force loop for the interparticle interactions in large (thousand or even million) particle simulations. This degree of possible "hyper-parallelism" suggests a very good prognosis for highly scalable, high performance PIMD codes. Interestingly enough, such an effort has never been carried out to our knowledge. It is estimated that for the simplest systems PIMD code can be developed which achieves 100 Gflop performance on a large parallel computer.

In a non-trivial theoretical advance [2-5], PIMD methods have been extended to include quantum dynamical effects (i.e., not just equilibrium properties). This computational approach is called "Centroid Molecular Dynamics" (CMD). It incorporates the dominant quantum dynamical effects of a many-body system into a classical-like MD framework, thus significantly extending the range and applicability of MD methods. The basic result of CMD [2-5] is that the dynamical correlations of quantum particles in a general many-body system can be accurately computed by running classical-like trajectories on an effective potential which transparently includes the effects of quantum zero point energy and tunneling. The task of integrating the CMD equations is not trivial since the effective potential is actually a kind of quantum potential of mean force, requiring an "on the fly" dynamical averaging at each timestep. A number of powerful and flexible algorithms for solving the basic CMD equation have been developed [5]. The best CMD algorithm so far is one in which the natural CMD timestep (similar to the classical MD timestep) is broken into a number of smaller timesteps. At each of the configurations for these timesteps, several constrained MD simulations with appropriate thermostatting are carried out in parallel to thermally average the effective force over a fixed number of secondary averaging timesteps. The system variables are then moved with the effective force according to the CMD equations using the smaller timestep. This algorithm can always be carried out in parallel with the CMD integration being the "master" and the effective force averaging MD runs being the "slaves" [6]. In fact, a second tier of parallelism can be introduced in which, for each of the averaging slaves, the MD steps are sent to additional nodes for each discretized path integral imaginary time slice [6]. A third tier of parallelism can even be introduced for very large system where, for each imaginary time slice, the

classical force loop is split over several nodes. A useful by-product of the CMD approach is an efficient algorithm to find potential energy minima of complex, many-body systems.

The PIMD and CMD algorithms are of general interest and considerable utility in all areas of condensed matter computer simulation. However, they are particularly important for the Air Force High Energy Density Matter (HEDM) Program in the computational study of, e.g., solid hydrogen, helium, or nitrogen matrices which contain isolated reactive species such as lithium, boron, hydrogen atoms, organic radicals, etc. Such matrices are high priority targets as possible HEDMs for Air Force space propulsion purposes. At the temperatures appropriate to the condensed phases of hydrogen or helium, the host matrix molecules will exhibit large quantum mechanical effects, but the treatment of these effects by a direct attack on the time-dependent Schrödinger equation is impossible. The key issue is the stability of the HEDM (i.e., the ability to trap the energetic impurities for some period of time and thus impede their diffusion and eventual recombination). Both the structural and dynamical aspects of the composite condensed phase HEDM systems are of interest. For the equilibrium structural studies, PIMD methods are adequate. For the dynamical studies, the CMD approach must be employed to directly simulate the diffusive and recombination steps of the guest atoms in the solid host. The long-term goal of this research was to develop general computational methods to rapidly and efficiently characterize proposed high energy density materials. The P.I. is also supported by AFOSR to carry out scientific HEDM modeling studies, in close contact with DoD researchers at the AF Phillips Lab.

#### Accomplishments/New Findings:

Parallel versions of the PIMD and CMD codes for homogeneous systems such as cryogenic solid hydrogen HEDM were completed. A code for multi-site homogeneous molecular systems was also finished. Also, the code for multi-site inhomogeneous systems is finished. The parallel scaling of the codes has been established on the IBM SP and SGI O2000 systems, as well as Linux clusters (see below).

#### *PIMD Code for Homogeneous Atomic and Multi-site Systems*

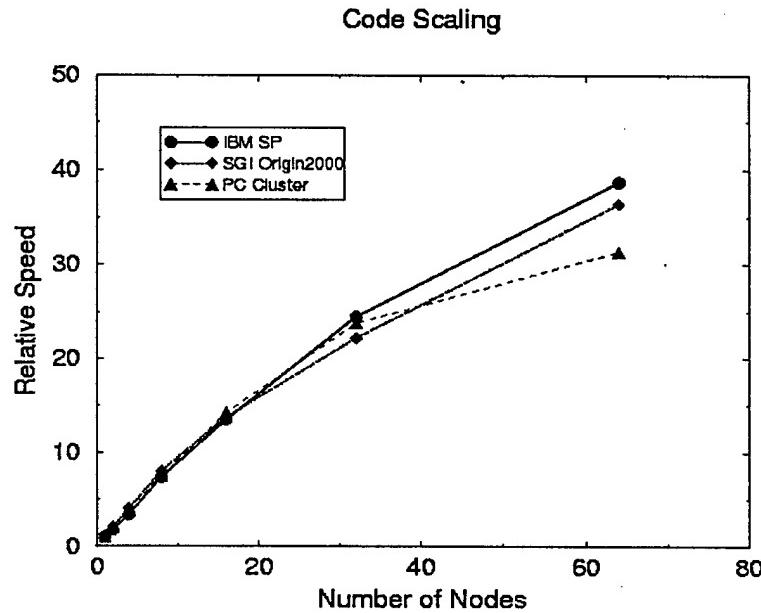
The code allows PIMD simulations of systems such as solid para-hydrogen with an arbitrary number of impurities (e.g., Li or B) and for molecular liquids or solids. Currently it can calculate the equilibrium structure and pair distributions. Also, it will give several thermodynamic quantities including energies. The code has constant temperature and/or constant pressure capabilities which can be turned on or off depending on the simulation requirements. Nose-Hoover chain dynamics are employed to maintain the system at the desired temperature if needed. Each quasiparticle has independent Nose-Hoover chains for the three different orthogonal directions with an arbitrary chain length to insure efficient canonical distribution of the system. Usually a chain with chain length four is more than good enough for canonical sampling. Increasing chain length more than four will only slow down the overall performance. The system can be kept under constant pressure by employing Anderson's extended system method where the volume of the system, i.e. simulation box size, is regarded as a dynamic variable. The three orthogonal simulation box lengths have independent scaling to better adapt to the anisotropic nature of solid para-hydrogen. The simulation parameters are adjusted to ensure fast convergence. The velocity Verlet algorithm is used as the numerical propagation scheme. Many of the parameters needed for simulations are adjustable depending on user level.

Force parallelization has also been implemented for even more calculation efficiency. This is useful if a large number of nodes are available, say more than 100 nodes. For force parallelization, we have used method by D. Walker which is a variation of the RD (replicated data)

algorithm. The number of total particles needs to be an integer times a force parallelization factor if this feature is used. The force parameter input formats are AMBER compatible.

### *CMD Code for Homogeneous Atomic and Multi-site Systems*

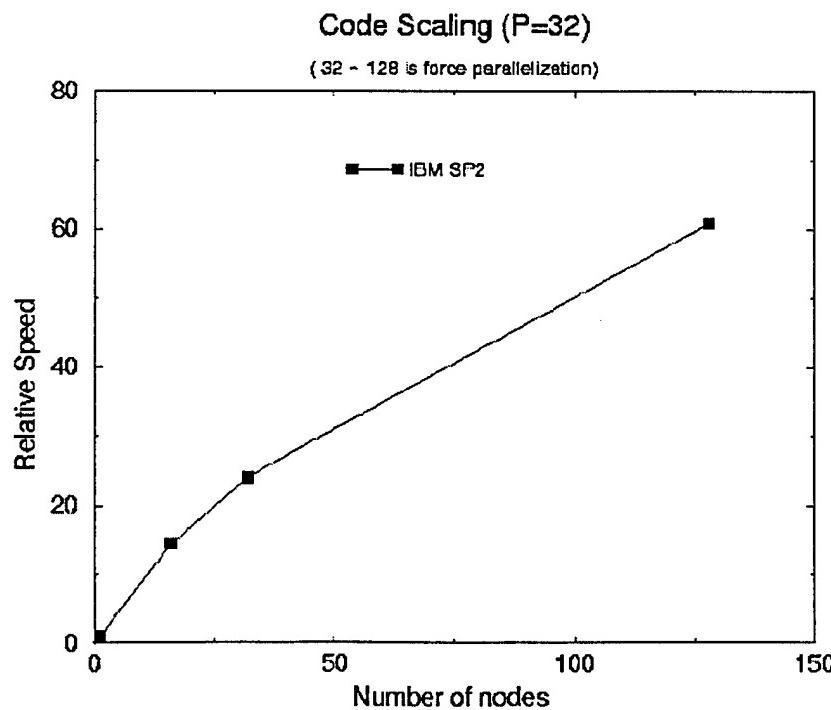
This code calculates constant volume CMD for systems such as solid para-hydrogen with an arbitrary number of impurities (e.g., Li or B) and for molecular liquids or solids such as water. Normal mode PIMD algorithms are used to speed up the convergence of the centroid force in the simulation. During the normal mode transformation, one of the coordinates naturally appears as the centroid coordinate. By fixing this coordinate at a given time, the remaining coordinates are dynamically propagated to get the average force acting on the centroid coordinate. The dynamic equation of motion of the centroid is governed by the average centroid force. The same Nose-Hoover chain dynamics used in the PIMD code are also used in the centroid force sampling. Also, the centroid coordinate has a Nose-Hoover chain which can be turned off at given time so that the dynamic behavior of centroid can be investigated after equilibration. Like the PIMD code, Velocity Verlet is used to propagate the system. For equilibrium properties, the code could calculate  $g(r)$  and energies also.



The code is hyper-parallelized in the sense that it is parallelized over quasiparticles and over independent centroid force averaging trajectories so that the total centroid force is summed over quasi-particles and averaged over independent trajectories for an efficient calculation of centroid forces. Usually, the starting point of this code is the result file from the above PIMD code, but it can also start from a lattice configuration, e.g., a hexagonal close packed structure. A plot of the scaling of the code on the IBM SP and SGI O2000 machines, as well as a Pentium cluster, is shown above. The force parallelization of the long range forces (Ewald summation) was achieved with the method of W. Smith. The code scales roughly as  $N^{0.9}$ , where  $N$  is the number of processors.

### *PIMD/CMD Code for Inhomogeneous Multi-Site Molecular Systems*

As an extension of the codes described above we have developed separate code for *Inhomogeneous Multi-Site Molecular Systems*. Our objective is the capability to simulate solute-solvent systems with arbitrary intramolecular connectivity and the number of solvent/solute is adjustable. Water is one of the important solvent in many cases. As a results, we have implemented quantum water (SPC/F2 model) as an option so that one might use water as default solvent in their simulation system if desired. The present code does both PIMD and CMD simulations of general molecules in a constant volume environment. Again, the code can calculate energies and  $g(r)$  between solvent-solute interactions by PIMD simulation and input format is AMBER compatible. Optional constant temperature sampling is obtained by Nose-Hoover chain dynamics. The topological structure and force parameters of molecules are treated as input files. Harmonic intramolecular forces are assumed. The code handles long range Coulombic forces between charged species by Ewald summation with fixed charge values once the partial charge of atom at each site is specified. The short range intermolecular non-bonded interactions have a 12-6 type Van der Waals potential. The number of quantum quasiparticles can be changed, and these can have different values depending on the atomic site of the molecule. This feature is so that some light atoms can have many quasiparticles (such as hydrogen) compared other heavy atoms (such as oxygen) which have fewer or are treated classically. For the PIMD code, it can calculate various thermodynamic averages for the system and it will generate equilibrium configurations. The CMD calculation can thus be started from the PIMD configuration files. Like the CMD code for homogenous single-site atomic systems, this CMD code has an on/off thermostat switch for the centroid variables. Both the PIMD and CMD versions of the code use path integral normal mode coordinate transformations to ensure fast convergence of the relevant averages. The code is parallelized over quantum quasi-particles as well as forces, with a scaling over quasiparticles similar to the code described earlier. The scaling of the code over quasiparticles and forces is shown below, up to 128 processors. The scaling is first over 32 processors for 32 quasiparticles, then over 4 processors for the force calculations for each quasiparticle.



## **REFERENCES:**

- [1] See, e.g., B. J. Berne and D. Thirumalai, Annu. Rev. Phys. Chem. **37**, 401, (1987).
- [2] J. Cao and G. A. Voth, J. Chem. Phys. **99**, 10070 (1993).
- [3] J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5106 (1994).
- [4] J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6157 (1994).
- [5] J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6168 (1994).
- [6] A. Calhoun, M. Pavese, and G. A. Voth, Chem. Phys. Lett. **262**, 415 (1996).

## **Personnel Supported:**

Gregory A. Voth (P.I.)  
Soonmin Jang (Postdoc)  
Srinivasan Iyengar (Postdoc)  
Jielou Liao (Postdoc)  
Udo Schmitt (Postdoc)  
Pierre-Nicholas Roy (Postdoc)  
Aaron Lefohn (Graduate Student)

## **Publications:**

This project is primarily a software development effort for the DoD high performance computing community. Publications were not a primary objective of this effort. However, one publication was: M. Pavese, S. Jang, and G. A. Voth, "Centroid Molecular Dynamics: A Quantum Dynamics Method Suitable for the Parallel Computer," Parallel Computing **26**, 1025-1041 (2000).

## **Interactions/Transitions:**

### **(a) Participation/Presentations at Meetings, Conferences, Seminars, etc.**

None related to this software development effort.

### **(b) Consulting/Advisory Functions**

Gaussian, Inc. (Academic Partner)

### **(c) Transitions**

The parallel PIMD/CMD computational methodology has been transitioned to all DoD High Performance Computer Center users through the DoD CHSSI project (Contact: Dr. Ruth Pachter, Wright-Patterson AFB).

## **New Discoveries, Inventions, Patents**

None

## **Honors/Awards:**

Elected as Fellow of the American Association for the Advancement of Science (AAAS) and American Physical Society (APS); Reilly Lecturer (University of Notre Dame); Frontiers of Chemistry Lecturer (Wayne State University); IBM Corporation Faculty Research Award